Principal Process Analysis of biological models

Stefano Casagranda, Delphine Ropers, Jean-Luc Gouzé

Journées annuelles du GT Bioss, Montpellier

Context and Objective

- Mathematical models of biological systems of high dimension

- Dynamics of large models are difficult to analyze:

• *Which regulatory mechanisms are important for the system dynamics?*

• *Do they always play a role during the dynamics?*

- Need to develop mathematical methods to answer these questions

• *Simplify the mathematical structure of the model*

• *Study the variation of activity of the remaining processes during the dynamics*

- Applied on Ordinary Differential Equation Systems

$$\dot{x}_i = \sum_j f_{ij}(x, p)$$

Values of parameters and initial values are known

Model Reduction Approaches

-Sensitivity Analysis

-Quasi Steady State Approximation

-Piece-wise affine differential equations

Other Similar Approaches

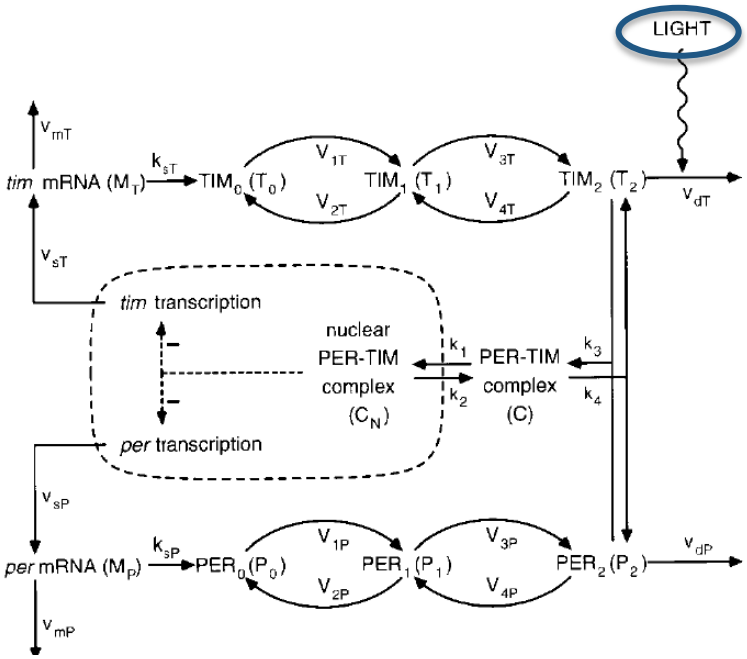
L. Petzold and W. Zhu, "Model reduction for chemical kinetics: An optimization approach," AIChE Journal, vol. 45, no. 4, pp. 869–886, 1999.

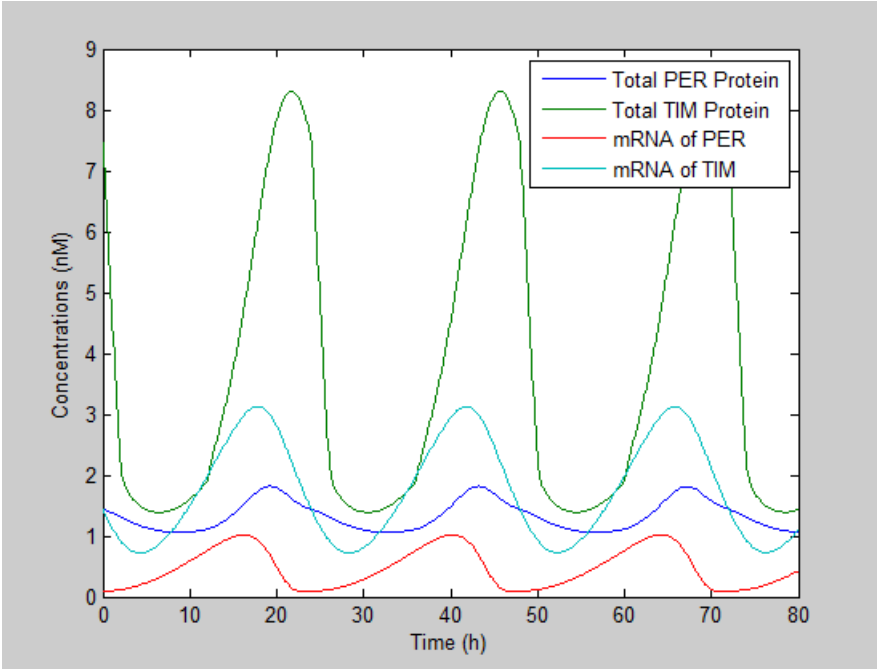
M. Apri, M. de Gee, and J. Molenaar, "Complexity reduction preserving dynamical behavior of biochemical networks," Journal of theoretical biology, vol. 304, pp. 16–26, 2012.

Circadian Clock

It allows the organisms to coordinate their physiological behavior with daily and seasonal changes in the day-night cycle (biological clock)

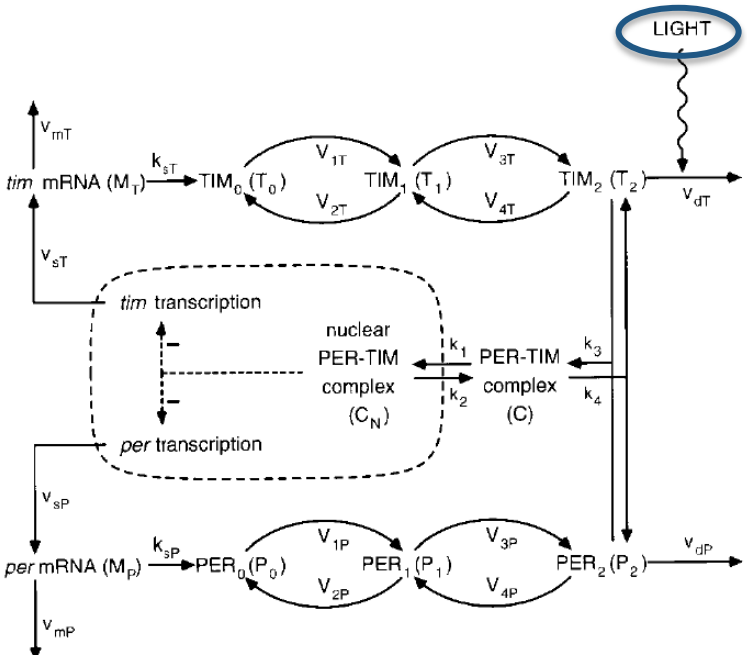
Model for circadian oscillations in *Drosophila* involving negative regulation of gene expression by PER and TIM gene





Leloup and Goldbeter (1998), *J Biol Rhythms*, 13(1):70-87

The model



ODEs

$$\frac{dM_p}{dt} = v_{sP}\frac{K_{IP}^n}{K_{IP}^n + C_N^n} - v_{mP}\frac{M_P}{K_{mP}+M_P} - k_d M_P$$

$$\frac{dP_0}{dt} = k_{sP}M_P - V_{1P}\frac{P_0}{K_{1P}+P_0} + V_{2P}\frac{P_1}{K_{2P}+P_1} - k_d P_0$$

$$\frac{dP_1}{dt} = V_{1P}\frac{P_0}{K_{1P}+P_0} - V_{2P}\frac{P_1}{K_{2P}+P_1} - V_{3P}\frac{P_1}{K_{3P}+P_1} + V_{4P}\frac{P_2}{K_{4P}+P_2} - k_d P_1$$

$$\frac{dP_2}{dt} = V_{3P}\frac{P_1}{K_{3P}+P_1} - V_{4P}\frac{P_2}{K_{4P}+P_2} - k_3 P_2 T_2 + k_4 C \quad - V_{dP}\frac{P_2}{K_{dP}+P_2} - k_d P_2$$

$$\frac{dM_T}{dt} = v_{sT}\frac{K_{IT}^n}{K_{IT}^n + C_N^n} - v_{mT}\frac{M_T}{K_{mT}+M_T} - k_d M_T$$

$$\frac{dT_0}{dt} = k_{sT}M_T - V_{1T}\frac{T_0}{K_{1T}+T_0} + V_{2T}\frac{T_1}{K_{2T}+T_1} - k_d T_0$$

$$\frac{dT_1}{dt} = V_{1T}\frac{T_0}{K_{1T}+T_0} - V_{2T}\frac{T_1}{K_{2T}+T_1} - V_{3T}\frac{T_1}{K_{3T}+T_1} + V_{4T}\frac{T_2}{K_{4T}+T_2} - k_d T_1$$

$$\frac{dT_2}{dt} = V_{3T}\frac{T_1}{K_{3T}+T_1} - V_{4P}\frac{T_2}{K_{4T}+T_2} - k_3 P_2 T_2 + k_4 C \quad - V_{dT}\frac{T_2}{K_{dT}+T_2} - k_d T_2$$

$$\frac{dC}{dt} = k_3 P_2 T_2 - k_4 C \quad - k_1 C \quad + k_2 C_N$$

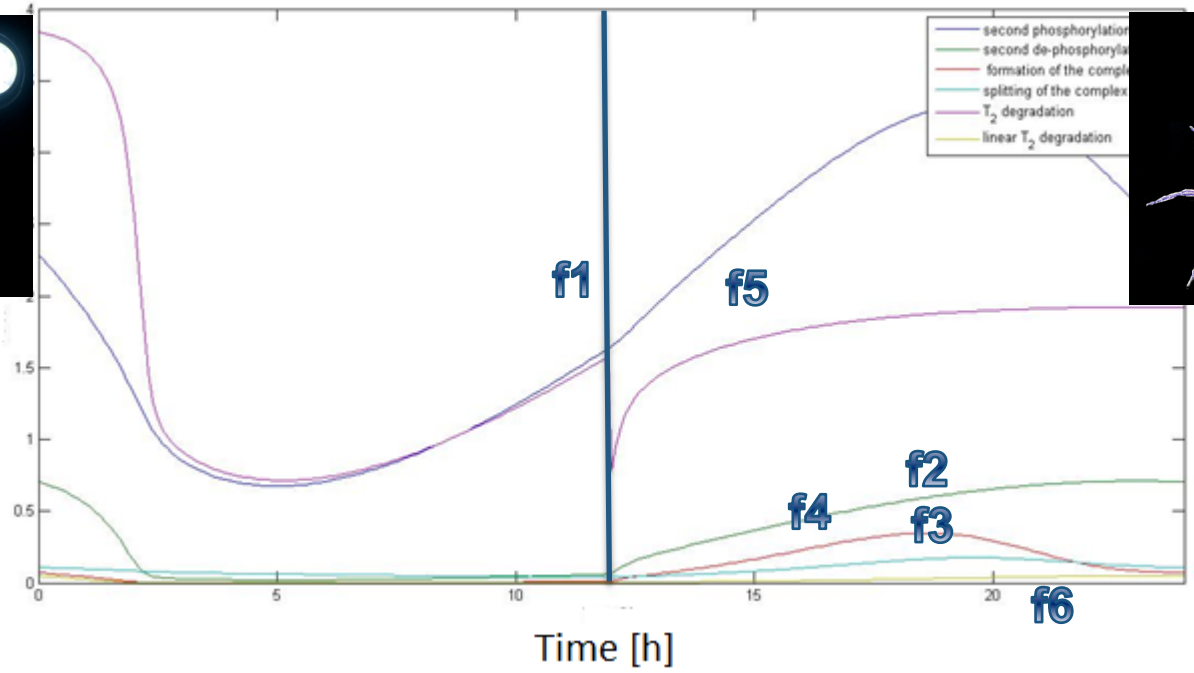
$$\frac{dC_N}{dt} = k_1 C - k_2 C_N$$

Ideas

Simulate the different processes for each ODE

$$\dot{x}_i = \sum_j f_{ij}(x, p)$$

$$\frac{dT_2}{dt} = \underbrace{V_{3T}\frac{T_1}{K_{3T}+T_1}}_{f1\ +} - \underbrace{V_{4T}\frac{T_2}{K_{4T}+T_2}}_{f2\ -} - \underbrace{k_3 P_2 T_2}_{f3\ -} + \underbrace{k_4 C}_{f4\ +} - \underbrace{\nu_{dT}\frac{T_2}{K_{dT}+T_2}}_{f5\ -} - \underbrace{k_d T_2}_{f6\ -}$$

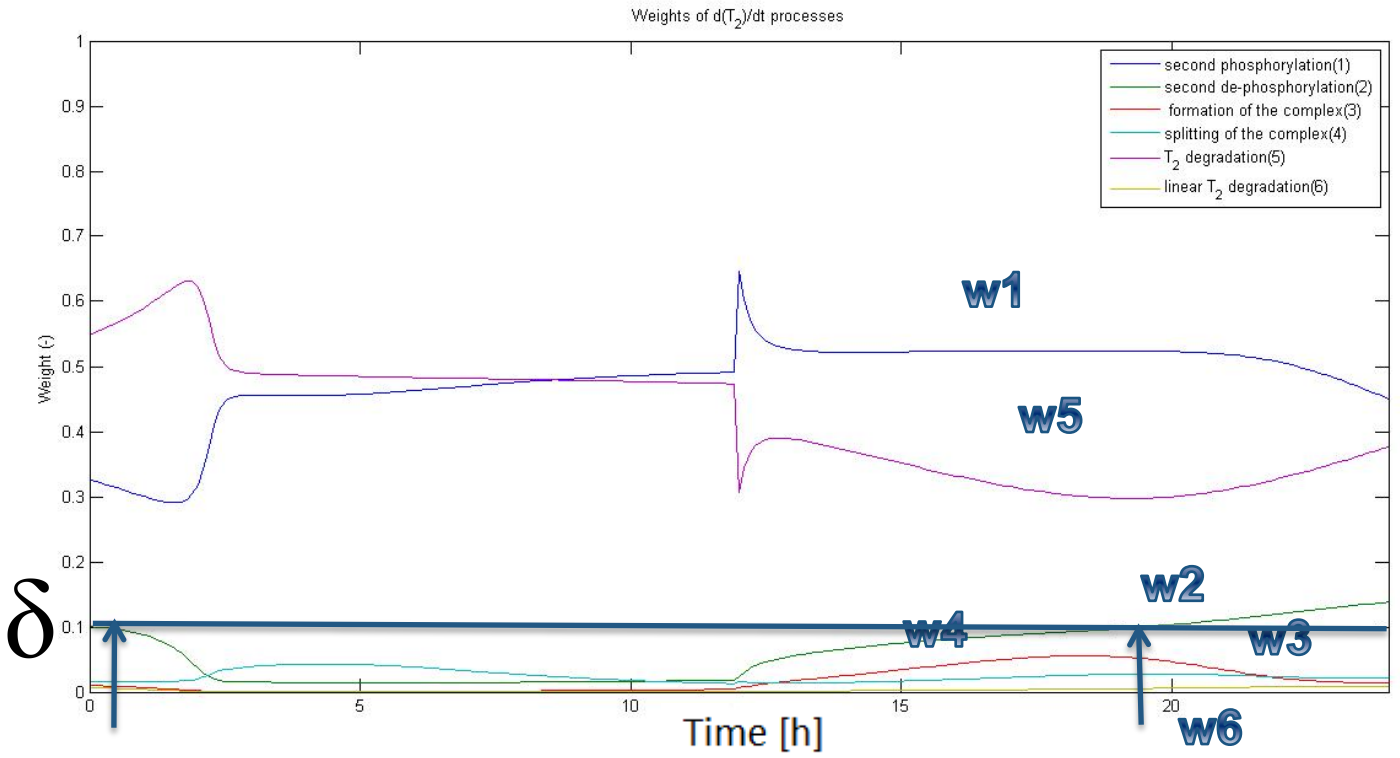


Ideas

Associate a dynamic relative weight for each process

$$W_{ij}(t,p) = \frac{|f_{ij}(x(t),p)|}{\sum_j |f_{ij}(x(t),p)|}$$

Ex: w1=f1/(f1+f2+f3+f4+f5+f6)



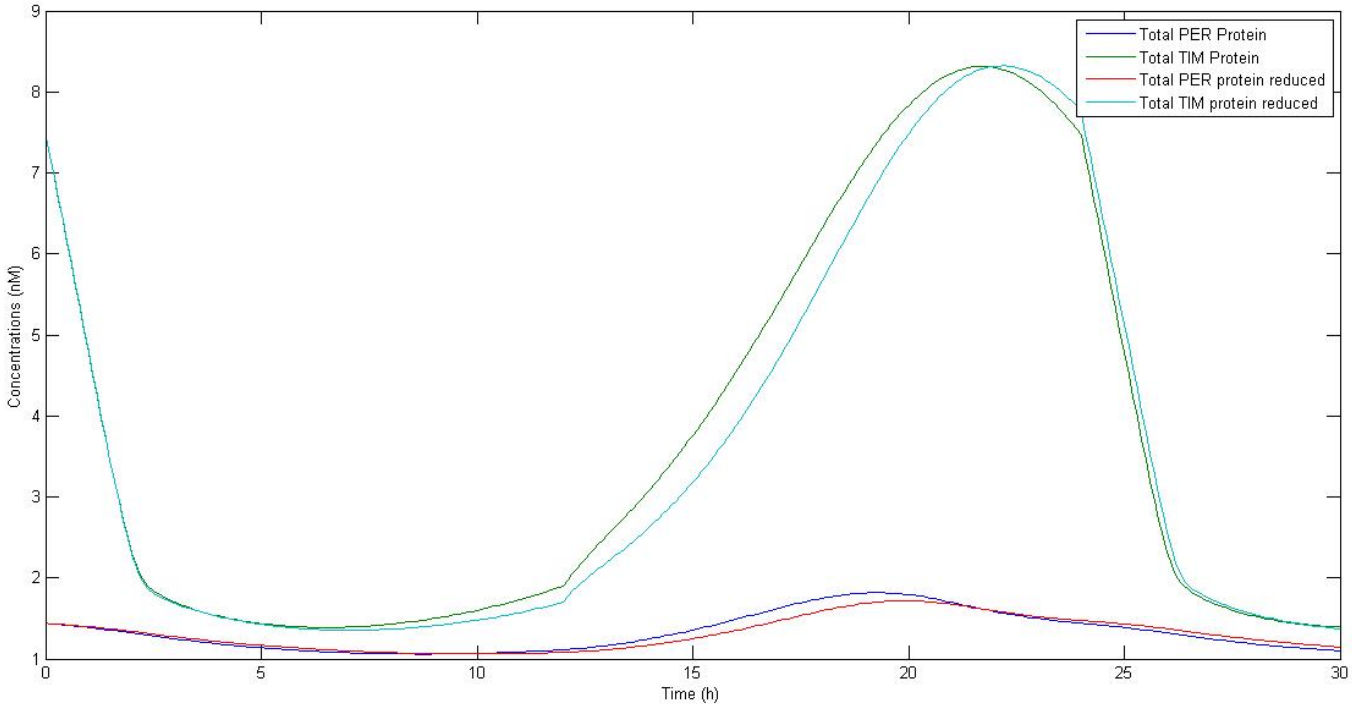
First step

Show how important processes evolve over time and when they can be considered “active”

Processes	0	t1	t2	t3	t4	t5	t6	t7	t8	t9	t10
Phosphorylated Timeless Protein											
First Phosphorylation	■	■	■	■	■	■	■	■	■	■	■
First De-Phosphorylation											
Second Phosphorylation	■	■	■	■	■						
Second De-Phosphorylation											■
Linear Timeless Period degradation											
Double Phosphorylated Timeless Protein											
Second Phosphorylation	■	■	■	■	■	■	■	■	■	■	■
Second De-Phosphorylation	■	■									■
Formation of the Complex											
Splitting of the Complex											
Double Phosphorylated Timeless degradation	■	■	■	■	■	■	■	■	■	■	■
Linear Double Phosphorylated Timeless degradation											
Complex											
Formation of the Complex	■	■					■	■	■	■	■
Splitting of the Complex	■	■	■	■	■	■	■	■	■	■	■
Shift in the Nucleus	■	■	■	■	■	■	■	■	■	■	■
Shift out the Nucleus	■	■	■	■	■	■	■	■	■	■	■
Linear Complex degradation											
Nuclear Complex											
Shift in the Nucleus	■	■	■	■	■	■	■	■	■	■	■
Shift out the Nucleus	■	■	■	■	■	■	■	■	■	■	■
Linear Nuclear Complex degradation											

First step

Simplify model by eliminating processes that are always negligible

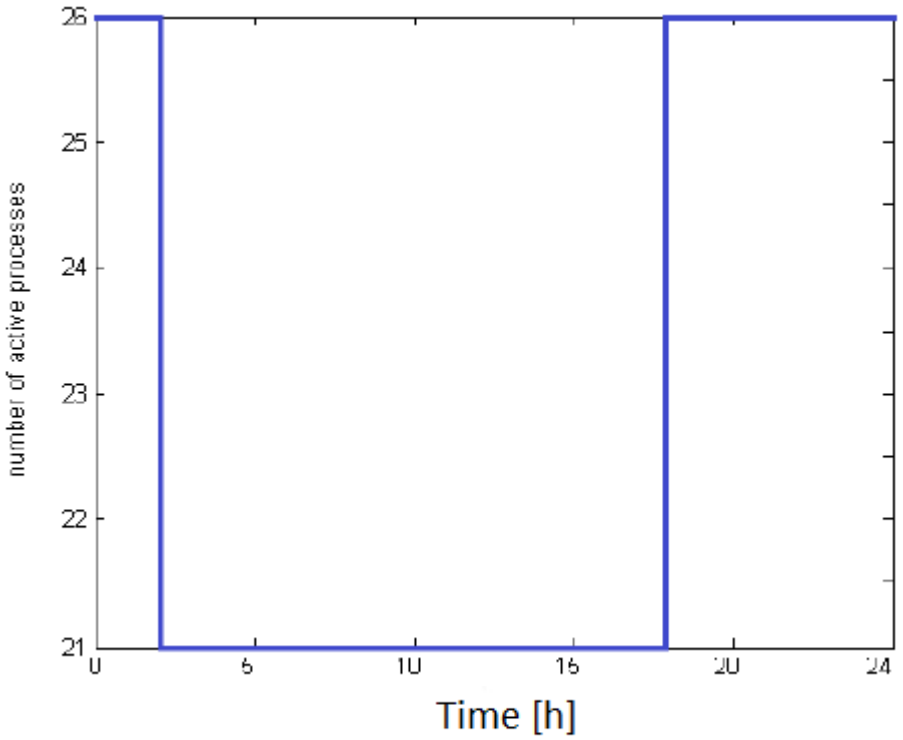


Second step

Create a “based-event” grid based on switching times and reduce it using clustering technique

Whitin-Cluster Sum of Squares

$$argmin_C \sum_{i=1}^{k} \sum_{x \epsilon C_i} ||x - \mu_i||^2$$



Second step

Create a chain of sub-models based on compacted time windows

From 0 to 1.96 h and from 17.8 to 24 h

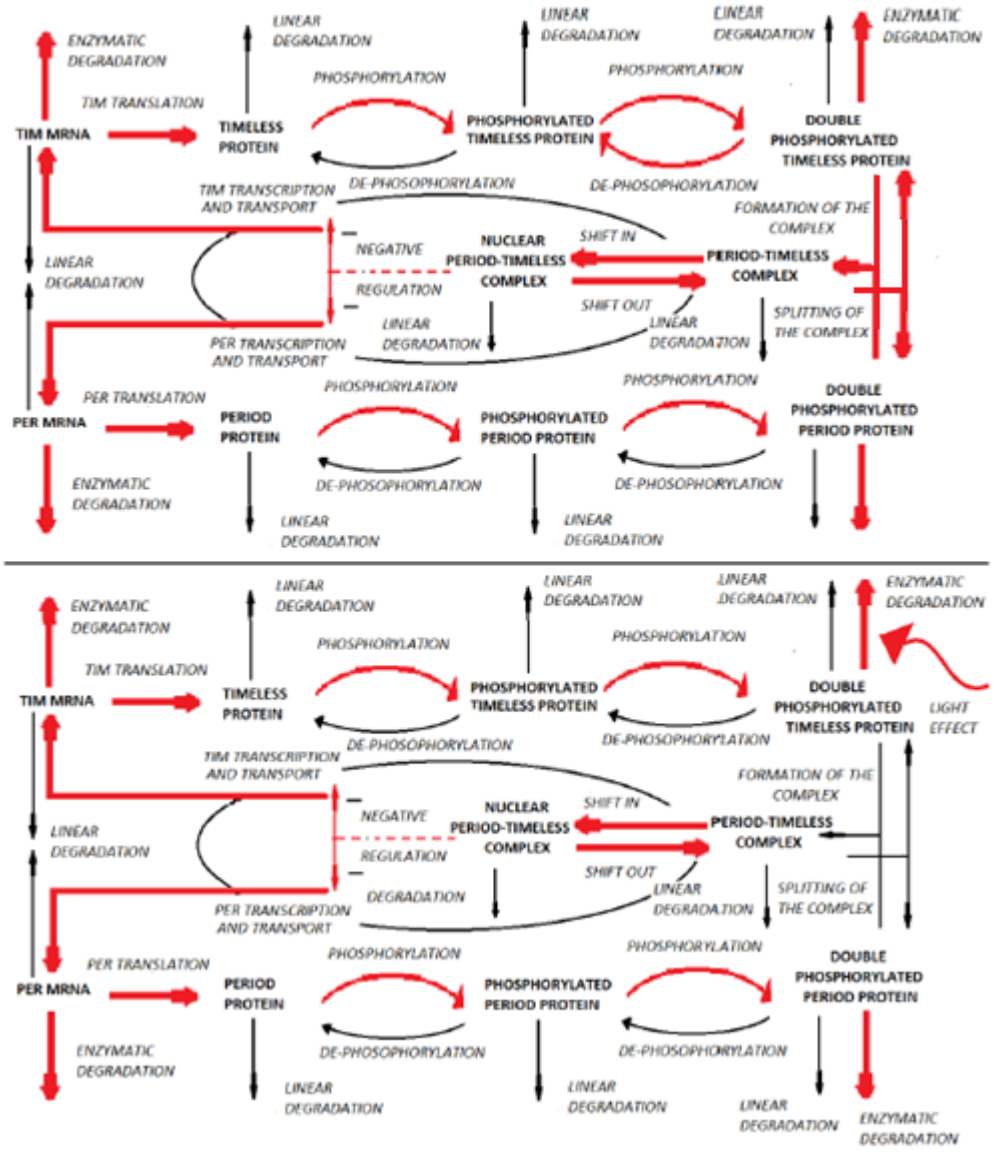
Variable	G. Rel. Err. S1 (%)
Period mRNA	13.63
Total Period Protein	1.61
Timeless mRNA	9.95
Total Timeless Protein	2.64
Complex	4.74
Nuclear Complex	5.36

From 1.96 h to 17.8 h

Variable	G. Rel. Err. S2 (%)
Period mRNA	7.70
Total Period Protein	7.06
Timeless mRNA	5.96
Total Timeless Protein	10.96
Complex	3.97
Nuclear Complex	5.85

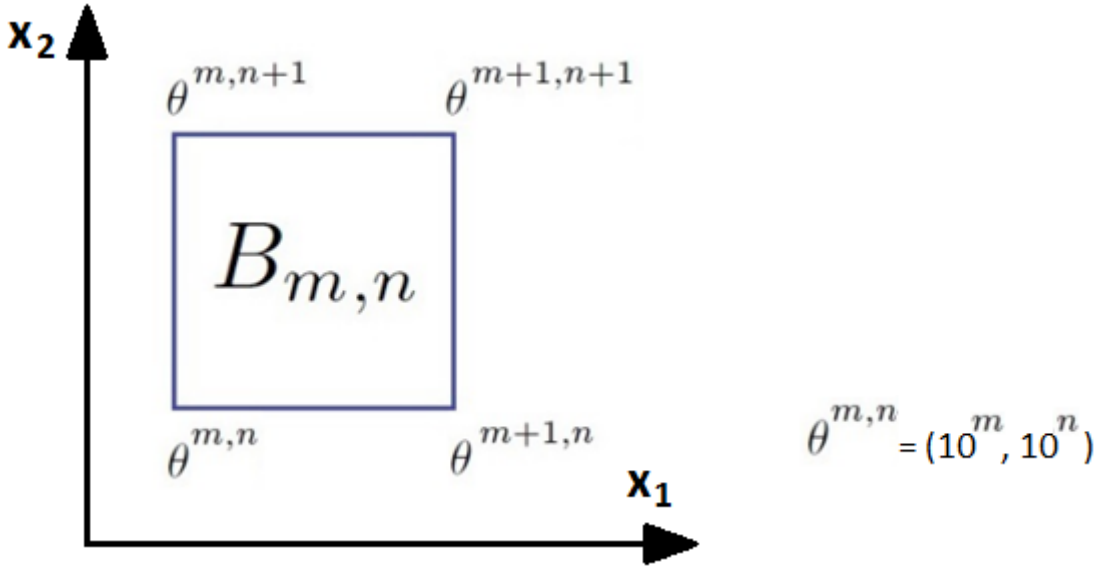
NIGHT TIME

DAY TIME



Process Analysis inside a rectangle

-Study effect of initial values on the outcome of reduced models

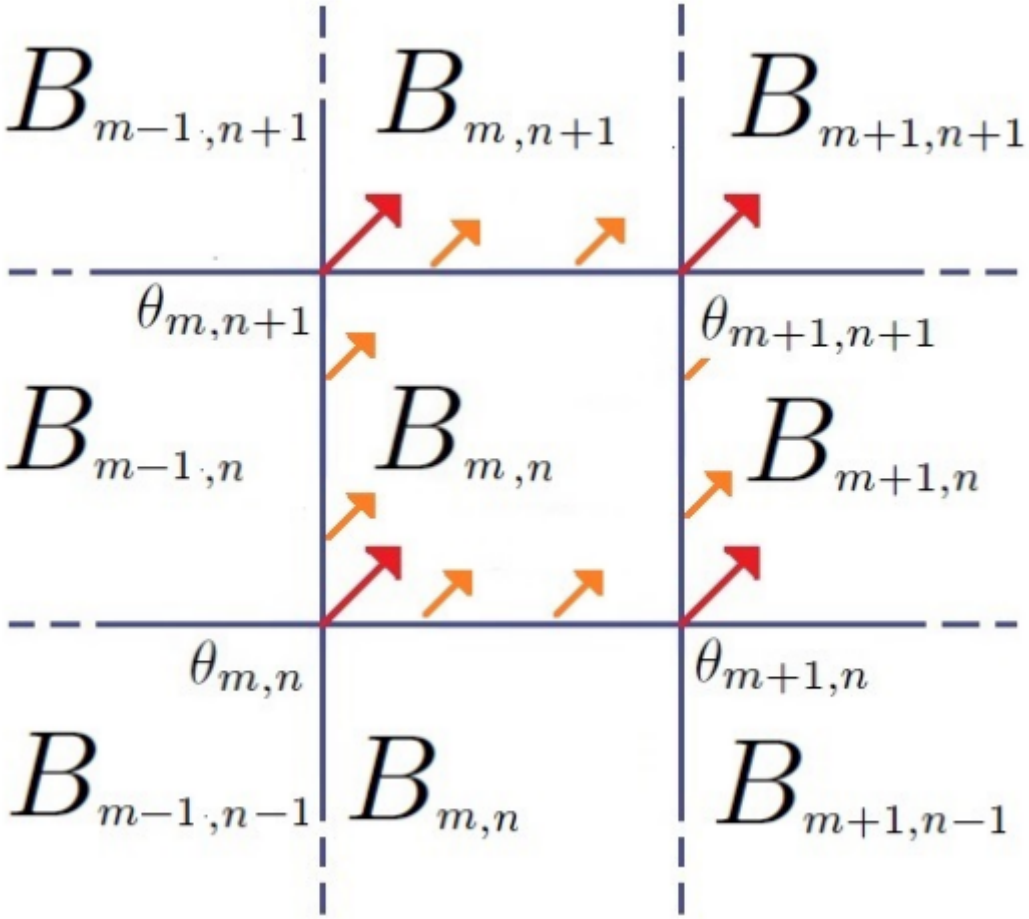


Assumption: *The Jacobian matrix* $J = Df(x,p)$ *of the system has a fixed sign inside the rectangle* $B_{m,n}$

- Neglect inactive processes inside every rectangle

$$W_{i,j}^{B_{m,n}}(p) = \frac{|f_{i,j}(S_{i,j}^{m,n}, p)|}{\sum_j |f_{i,j}(S_{i,j}^{m,n}, p)|}$$

Possible transition between domains

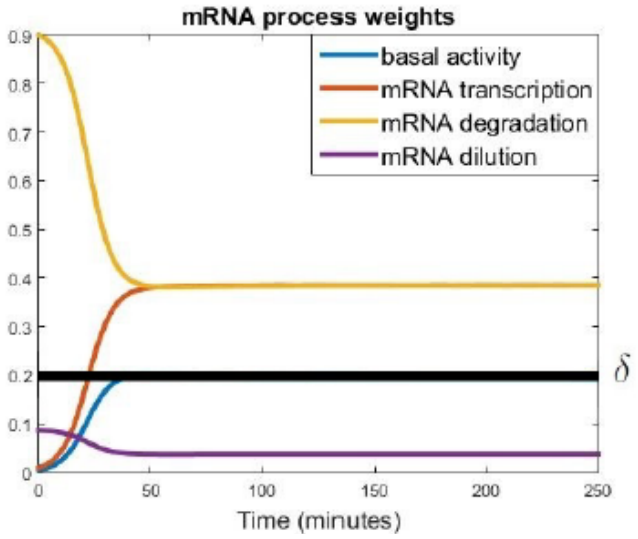


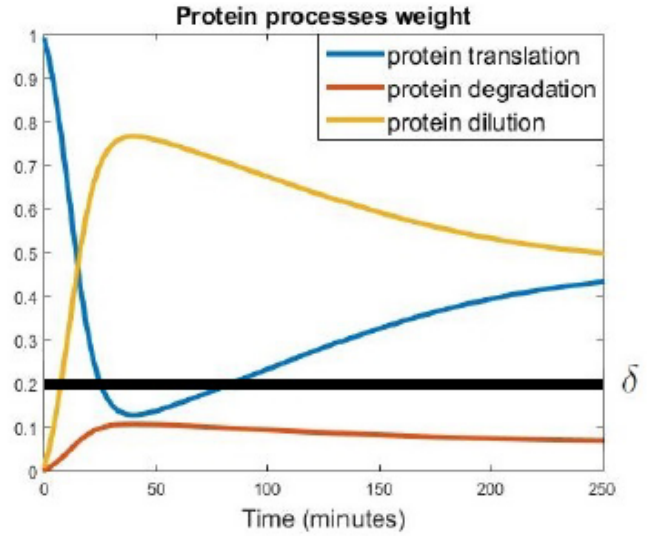
Gene expression model

$$\frac{d}{dt}M = \kappa_1 + \kappa_2 \frac{\alpha_P^m}{\alpha_P^m + P^m} - \gamma_M M + \mu M$$

$$\frac{d}{dt}P = \kappa_3 M - \gamma_P P + \mu P$$

$$\theta^{0,0}$$

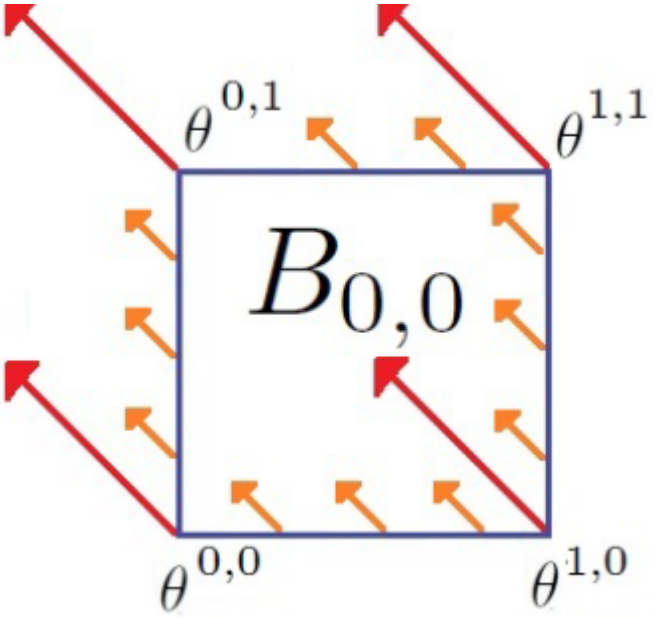


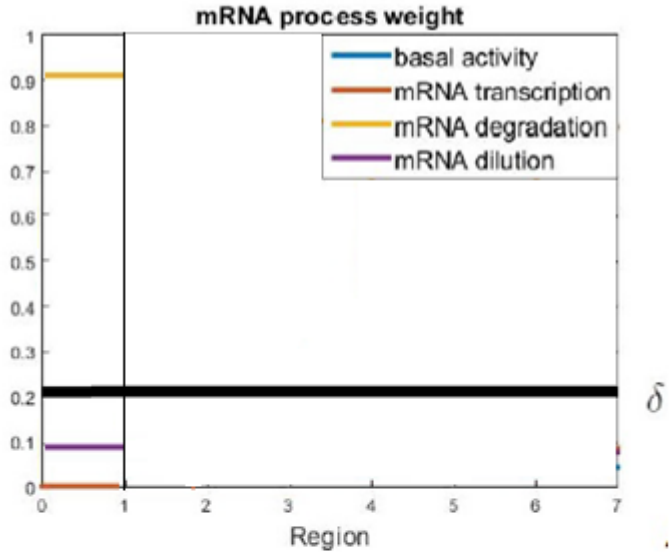


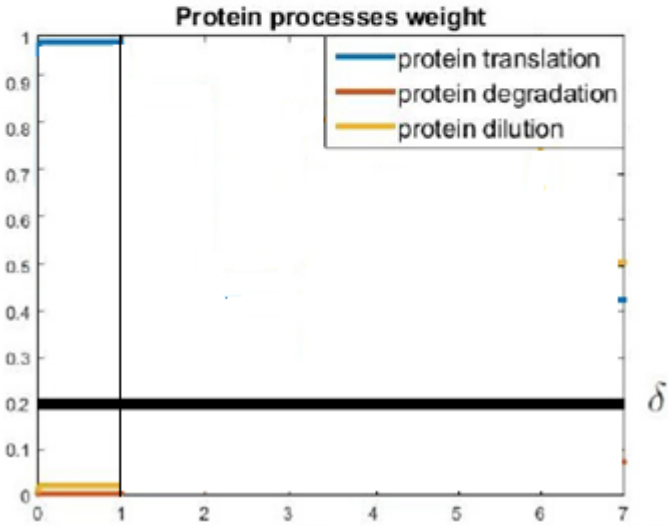
Gene expression model

$$\frac{d}{dt}M = \kappa_1 + \kappa_2 \frac{\alpha_P^m}{\alpha_P^m + P^m} - \gamma_M M + \mu M$$

$$\frac{d}{dt}P = \kappa_3 M - \gamma_P P + \mu P$$



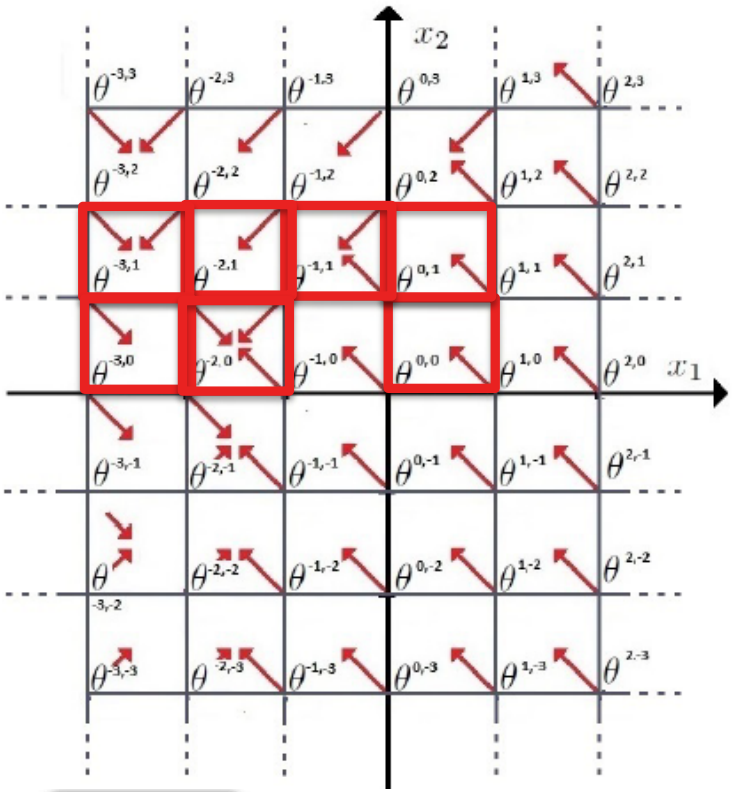


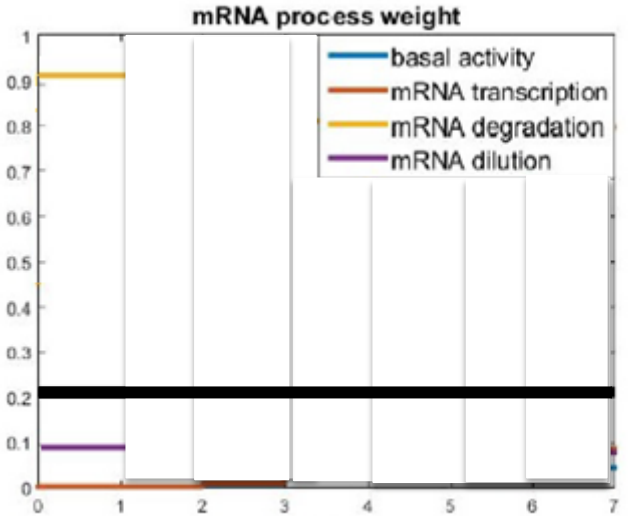


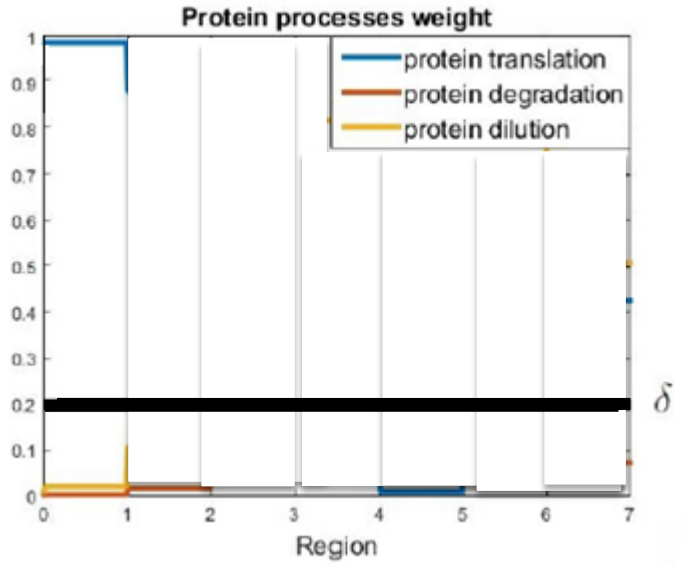
Gene expression model

$$\frac{d}{dt}M = \kappa_1 + \kappa_2 \frac{\alpha_P^m}{\alpha_P^m + P^m} - \gamma_M M + \mu M$$

$$\frac{d}{dt}P = \kappa_3 M - \gamma_P P + \mu P$$







Gene expression model

	MRNA PROCESSES				PROTEIN PROCESSES		
RECTANGLE	BASAL ACTIVITY	TRANSCRIPTION	DEGRADATION	DILUTION	TRANSLATION	DEGRADATION	DILUTION
B(-3,-3)		■	■		■		
B(-3,-2)		■	■		■		
B(-3,-1)		■	■		■		■
B(-3,0)		■	■				■
B(-3,1)		■	■				■
B(-3,2)	■		■				■
B(-3,3)	■		■				■
B(-2,-3)			■		■		
B(-2,-2)			■		■		
B(-2,-1)			■		■		
B(-2,0)			■				■
B(-2,1)			■				■
B(-2,2)			■				■
B(-2,3)			■				■
B(-1,-3)			■		■		
B(-1,-2)			■		■		
B(-1,-1)			■		■		
B(-1,0)			■		■		
B(-1,1)			■		■		■
B(-1,2)			■				■
B(-1,3)			■				■
B(0,-3)			■		■		
B(0,-2)			■		■		
B(0,-1)			■		■		
B(0,0)			■		■		
B(0,1)			■		■		
B(0,2)			■		■		■
B(0,3)			■				■
B(1,-3)			■		■		
B(1,-2)			■		■		
B(1,-1)			■		■		
B(1,0)			■		■		
B(1,1)			■		■		
B(1,2)			■		■		
B(1,3)			■		■		■

Conclusion

-We developed a method to analyze the role of regulatory mechanisms in the system dynamics where we gained knowledge about which and when mechanisms are at work

-We created a simpler model in which negligible mechanisms are not included and we decompose it into a succession of sub-models containing the core mechanisms

-We studied the effect of initial values on the outcome of the reduced models and we studied the transitions between different space regions

-PPA is a simple-to-use method, which constitutes an additional and useful tool for analyzing the complex dynamical behavior of biological systems.

Current/Future steps

- We used global relative errors to assess the quality of the model reduction and apply global sensitivity analysis to test the influence of model parameters on the errors.

- We are studying a refinement of PPA by considering three different levels of activities (inactive, active, fully active), defined by two different thresholds in order to improve the quality of model analysis and reduction.

-We are studying how to apply PPA on the full coupled system of equations instead of working on each equation separately: this would help to analyze activities or inactivities of processes shared by several equations.

Applied on…

Drosophila circadian Rhythms and cellular signal models

S. Casagranda, D. Ropers, J.-L. Gouzé.
Model reduction and process analysis of biological models,
in: Control and Automation (MED), 2015 23rd Mediterranean Conference on, IEEE, 2015, pp. 1132–1139.

Simple Gene Expression model

S. Casagranda, J.-L. Gouzé,
Principal Process Analysis and reduction of biological models with order of magnitude,
in: The 20th IFAC world congress, 2017-accepted.

Mammalian circadian clock model

S. Casagranda, S.Touzeau, D.Rophers, J.-L. Gouzé
Principal Process Analysis of biological models,
Journal of Theroretical Biology, 2017-submitted

Toxicological model

S. Casagranda, Frédéric Dayan, , J.-L. Gouzé, David Rouquié (Bayer CropScience)
Principal Process Analysis applied to a model of endocrine toxicity induced by Fluopyram
Ongoing Paper

H. Pagel, C. Poll, J. Ingwersen, E. Kandeler, T. Streck,
Modeling coupled pesticide degradation and organic matter turnover: From gene abundance to process rates, Soil Biology and Biochemistry 103 (2016) 349-364.

Fed- Batch cultures model

C. Robles-Rodriguez, C. Bideaux, S. Guillouet, N. Gorret, G. Roux, 490 C. Molina-Jouve, C. Aceves-Lara,
Multi-objective particle swarm optimization (mopso) of lipid accumulation in fed-batch cultures, in:
Control and Automation (MED), 2016 24th Mediterranean Conference on, IEEE, 2016, pp. 979–984.

Thank you

Thanks to:

Conseil Régional PACA

Project Reset

Centre de recherche
Sophia Antipolis - Méditerranée

www.inria.fr/sophia

Current step

-We are applying Parameter Sensitivity Analysis to sub-models to test their robustness

$$e_i = \frac{\int |x_i(t) - x_i^r(t)|dt}{\int |x_i(t)|dt}$$

p_j^-, p_j, p_j^+ → Number of Levels

38 → Number of Parameters

3^{38} → Too many simulations!!!

FRACTIONAL FACTORIAL DESIGN

$$e_i = \mu + \sum_j \alpha_j + \sum_j \sum_{k \neq j} \beta_{jk} + \epsilon_i$$

TOTAL SENSITVITY INDEX FOR EACH PARAMETER

$$tSI_j^i = \frac{SS_j^i + \sum_{k \neq j} SS_{j,k}^i}{SS_T^i}$$

Future step

-Analysis of the Gene Expression Machinery Model of Delphine Ropers